Journal Pre-proof

Towards a more inclusive and equitable developmental cognitive neuroscience

Jazlyn Nketia, Dima Amso, Natalie Hiromi Brito
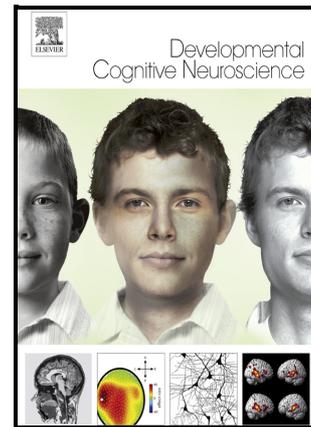
Please cite this article as: Jazlyn Nketia, Dima Amso and Natalie Hiromi Brito, Towards a more inclusive and equitable developmental cognitive neuroscience, *Developmental   Cognitive   Neuroscience,* (2021) doi:https://doi.org/10.1016/j.dcn.2021.101014

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Title: **Towards a more inclusive and equitable developmental cognitive neuroscience**

Jazlyn Nketia[a], Dima Amso[b], Natalie Hiromi Brito[c]

[a]M.S. Graduate Student Brown University Visiting Graduate Student Columbia University

[b]PhD Professor of Psychology Department of Psychology Columbia University

[c]PhD Assistant Professor of Developmental Psychology Department of Applied Psychology New York University

Abstract

Brain and cognitive development is a burgeoning area of scientific inquiry, with tremendous potential to better the lives of children. Large scale longitudinal neuroimaging studies offer opportunities for significant scientific advances in our understanding of developing brain structure and function. The proposed manuscript will focus on the scientific potential of the HEALthy Brain and Cognitive Development (HBCD) Study, highlighting what questions these data can and what they cannot answer about child development. Specifically, we caution against the misuse of these data for advancing de-contextualized and scientifically questionable narratives about the development of children from marginalized communities. We will focus on building and organizing a framework for interpreting HBCD data through the lens of sampling, cultural context, measurement, and developmental science theory. Our goal is to thoughtfully offer the scientific community opportunities to use the large scale and collaborative nature of HBCD to collectively revise practices in developmental science that to-date have not carefully considered their own role in perpetuating narratives that support systemic injustice.

Keywords:

HBCD; social justice; cultural context; brain; child development; bias

Introduction

Within the context of 2020-2021, from the racial disparities in the communities hit hardest by the COVID-19 pandemic, to the massive worldwide protests in support of the Black Lives Matter movement, there has been an increase in attention to the many systemic injustices within our American society. Although the experiences of the past year have focused most prominently on healthcare and the criminal justice system, introspection across multiple sectors of life, including academia, is warranted. In anticipation of the upcoming HEALthy Brain and Child Development (HBCD) study, we propose that this collaborative effort is an opportunity for scientists within developmental cognitive neuroscience to think deeply about the potential impact and consequences stemming from the interpretation of outcomes from this large-scale project.

The Healthy Brain and Child Development Initiative

The 21st century is a particularly exciting time for discovery and progress in the field of human development. In the past several decades, this area of study has rapidly built foundational knowledge

that informs principled questions about the drivers of change in human brain and cognitive development. This theoretical progress is paired with the development of multiple safe methods that can simultaneously probe genetic, brain, behavioral, and environmental factors beginning prenatally. Recognizing this unique convergence, the National Institute of Drug Abuse (NIDA) at the National Institutes of Health (NIH) has been working with world-class scientists across multiple sites in the United States to design a prospective examination- beginning prenatally and extending into childhood- of brain, cognitive, socioemotional, and behavioral development. Key to the outcomes of the HBCD study is an understanding of typical neurocognitive development, as well as clues as to how early prenatal and postnatal drug exposure and adversity may impact the long-term behavioral and mental health of developing children.

The potential power of this study cannot be understated. The goal is to enroll roughly 7500 pregnant women from multiple sites across the United States. A wide variety of influential factors, including prenatal and postnatal exposures to substances (opioids, alcohol), family stress, socioeconomic status, toxic metals, sleep, nutrition, supportive caregiving, and medical history will be assessed longitudinally and paired with measures of brain and cognitive development. More importantly, this national study will last 10 years, resulting in the largest and most thorough data collection of its kind documenting early life. HBCD will dovetail with an existing study that similarly examines development in ~11,500 children longitudinally from the age of 9 through late adolescence (Adolescent Brain Cognitive Development [ABCD] Study). While we focus this piece on the upcoming HBCD study, we expect that many of the issues discussed here for HBCD will also apply to the ABCD study. Garavan et al., (2018) is an excellent discussion of sampling and design procedures for ABCD and Simmons et al., (2021) recently described procedures for responsible use of ABCD data.

Although the HBCD study's broad design follows that of the ABCD study, the HBCD study is intended to inform how exposures early in life increase risk for negative long-term outcomes in psychological, behavioral, and neural developmental domains. By definition, this is an epidemiological observational cohort study. This study design is fairly common in medicine and public health (Carlson & Morrison, 2009; Thiese, 2014), but not common to the brain and cognitive science communities. As such, it is critical to define during the earlier stages of this large project exactly what this type of research design can and cannot tell us about human development. In the sections that follow, we will briefly describe the strengths and limitations of observational studies, as well as issues that must be considered when reporting data from such studies. We will then explore the social and economic context surrounding data collection, as well as the potential confounds and biases that must be anticipated when using the dataset to make inferences about child development in disparate communities of American children. We will also specifically discuss the potential misuse and misinterpretation of these data relevant to perpetuating stereotypes and narratives about the development of children from marginalized communities. Finally, we will propose a framework for interpreting HBCD data through the lens of conscious sampling, cultural context, measurement, and interdisciplinary developmental science theory.

Observational Cohort Studies: Strengths and Limitations

Large cohort observational studies are intended to determine correlations or associations between measured variables. In the case of the HBCD study, the variables are early life exposures (e.g., stress, adversity, substances etc.) and assessments tapping into physical health, brain maturation, and cognitive performance. While it is tempting to represent this study design in terms of a causal framework, with early exposures acting as causes and measures of brain and behavior interpreted as outcomes,

observational studies offer no such specificity. Complex correlational relationships are often riddled with confounding or missing variables. For example, the same genetic or environmental variables that predispose a mother to use a substance prenatally, may also separately shape her infant's postnatal development.

True experimental design involves random sampling, manipulation of the independent variable, and measuring the dependent variable as part of the experiment. The strength of the HBCD study is in identifying these correlational relationships in a large representative cohort. By doing so, the findings from HBCD can narrow the space of experimentation for testing hypotheses about the causal relationship between two associated variables. *Thus, HBCD data can offer strong evidence for risk and protective factors that then must be followed up and tested within an experimental design.* It is notable that many of the associations that will be found cannot be ethically experimentally tested in human subjects. For example, it is not ethical to expose a pregnant woman to a teratogen to examine long term neurodevelopmental impact on her child. As such, animal models become very valuable. In our view, the inability to design an experiment to test particular associations within human subjects neither elevates nor undermines HBCD findings, but rather has specific ramifications for the interpretation of such results, namely the risk of being misinterpreted to reflect causation.

Absolute and relative risk.

How observational cohort studies are conventionally reported and interpreted is of utmost importance. In reference to specific terminology, 'risk factors' are variables that predispose someone to a particular outcome, whereas 'protective factors' can prevent such an outcome in the presence of risk. It is important to note that risk can be relative or absolute. In medical epidemiological studies, relative risk is given as an odds ratio between a subsample of a population that have encountered an experience or exposure of interest and another that is otherwise matched on confounding variables but has not had this specific exposure. Absolute risk, on the other hand, is the risk of a particular outcome for any single child who has had a defined exposure. To illustrate the difference, a relative risk assessment would report that a child with early life teratogen exposure is three times as likely to develop a neurodevelopmental disorder than a child who does not experience these adversities. However, the absolute risk to any one child in the exposure group may still be very low, as this value would be expressed as the number of children with a diagnosed neurodevelopmental disorder divided by the total number of children in the exposure sample.

Within the psychological sciences, statistical methods often report group-based differences using parametric statistics (i.e., relying solely on statistical significance or a *p value*). Consequently, this enables the detection of small effects by simply increasing the overall sample size. The propensity to emphasize statistical significance can be problematic and misleading precisely because it is heavily influenced by sample size (Zuo et al., 2019). This is particularly an issue for HBCD, as the sample size is quite large. Therefore, minor, and potentially inconsequential effects may be statistically reliable. Another critical metric then is effect size, which indicates whether measured differences are practically significant and do not depend on sample size (Flora, 2020; Serdar et al., 2021; Sullivan & Feinn, 2012). In sum, to best understand the real-world impact of statistically significant differences between groups, it is critical to supplement statistically significant findings with effect size estimates and both relative and absolute risk scores when reporting and interpreting results.

Normative variability and theory building.

We offer a final word about what constitutes "outcome" variables in large epidemiological studies. Generally, these outcomes are clinically, socially, or economically relevant variables measured at some determined endpoint. These might include heart disease in middle adulthood or anxiety disorders in adolescence. Variables like cortical volume or attention span may be mediators between risk factors and outcomes but are difficult to evaluate as outcomes themselves unless they are pathologically outside the range of normative variability. Indeed, perhaps the greatest challenge for scientists in responsible and accurate interpretations of their data are in cases where results may link experiences and exposures with outcomes that fall within the range of normative variability. Absolute and relative risk make a great deal of sense when assessing finite or clinically relevant outcomes, such as development of disease or disorder. However, their value is unclear when examining normative variability along some cognitive or socioemotional dimension that itself is loosely tied to negative outcomes.

The primary reason for this is that human development is not a march toward a stable adult state. The best science to date indicates that human development is adaptation to a child's ecological niche on short timescales (Johnson et al., 2015; Werchan & Amso, 2017). Given limited resources, any organism is likely to invest in specific domains that ensure adaptation and survival, at the cost of others. Within normative variability, there is no good brain or bad brain. There is only human diversity. Moreover, the distribution of participants into sub-cohorts represented by experiences or exposures to alcohol, opioids, adversity, foster care, inadequate health care, household/community violence and so on is not random. Unquestionably due to systemic social injustices in the United States, these experiences and exposures are more prevalent in lower-income, Black and Brown, and immigrant communities (Burger & Gochfeld, 2011; French et al., 2020; Williams et al., 2016). As a result, these exposures or 'risk factors' of interest are substantially more likely to correlate with race, ethnicity, and income - thereby greatly enhancing the probability of finding correlational associations between variables within the observational cohort design. *Thus, these statistically significant associations may reinforce biased interpretations of the lived experiences of families from marginalized communities and further perpetuate the systems and structures that uphold social injustice.* Controlling for one variable (e.g., race) while looking at the predictive value of another highly correlated and interrelated variable (e.g., income) is arguably problematic. Although there are instances where mathematically conceivable results may occur, they most likely reflect de-contextualized interpretations that considerably lack validity and reliability. Moreover, these interpretations are typically blind to the broad sociocultural and policy-based lived experiences of children and families in the United States. Nonetheless, there is a tendency to interpret associations between experience and outcome as relevant to a fixed maturational process and through the lens of common scientific biases. In the next sections, we review common scientific biases that will need to be addressed within the HBCD study in order to avoid the pitfalls of biased interpretations and maximize the utility of this potentially invaluable data set.

Common scientific biases to avoid and paths to pursue

In the following section we outline common interrelated scientific biases in sampling, generalizability, normativity, and theory building (see Hruschka et al., 2018). We want to clarify at the outset that it is our firm belief that any bias in science is one that renders the science inaccurate and invalid. The biases we will discuss range from internal validity (measurement) to external validity threats (sampling, measurement, normativity). Therefore, while we frame this piece with social justice in mind, we remind the reader that the *quality of the science itself* is also at stake. Without addressing these biases, data can

be misinterpreted and misapplied, ultimately undermining the potential positive impact of the information obtained. We are not simply proposing that all studies include large representative samples, as there are many legitimate reasons to focus on limited or narrow samples. Instead, we encourage researchers to specifically identify who the results may or may not generalize to, think about the varied experiences of populations not included within the sample when engaging in theory building, and most importantly, broaden their views of humanity and cultural context to include ideas of how context could *adaptively* shape development.

Sampling Bias

The magnitude of the proposed HBCD study provides an opportunity to recruit a representative sample of American families and shift the field of developmental cognitive neuroscience towards a more equitable standard of research. Efforts to obtain this standard must not end with the sampling strategy. Historically, developmental science samples include a disproportionately high representation of White, affluent families sometimes described using the acronym W.E.I.R.D. (Western, Educated, Industrialized, Rich and Democratic). The homogeneity of past research has both contributed to and perpetuated scientific biases that shape the field. A review of the top five high-profile developmental journals identified sampling methods that contribute to the lack of representation, and interpretative limitations that accompany these sampling methods (Bornstein et al., 2013). Depending on the journal, one-quarter to over two-thirds of the articles included in the review did not adequately report race and ethnicity. In fact, 41.4% of the published articles omitted or insufficiently reported ethnic and racial information, including those that only reported that the samples were ''predominantly White'' or ''about half minority.''

Representative and random sampling is the hallmark of unbiased science with human subjects. We recommend reviewing Bonevski et al., (2014) for a thorough discussion of various efforts and successes in recruitment of what they call 'hard-to-reach' underrepresented groups. In brief, they suggest that representative sampling involves both a serious investment in time and resources and local research-community partnerships. There are numerous challenges for establishing such partnerships. There is a well-documented history of unethical scientific medical practices with Black and Brown bodies in the United States (see Scharf et al., 2010; Washington, 2006). This of course contributes to the degree of trust or mistrust of academic institutions and research. Moreover, there is often a complex relationship between academic institutions, themselves bastions of privilege, and the surrounding communities. Research sites must carefully consider these issues through the lens and voices of community members, before establishing solutions for increasing representative sampling. Partnerships will (a) foster genuine trust, (b) increase the effectiveness of science communication, (c) alleviate the specific hurdles (e.g., time and financial) to participation faced by community members, (d) commit to learning what these communities' *actual* needs are, and if applicable (d) offering services, within the realm of their developmental science expertise. Partnership is indeed bidirectional. In the sections that follow, we suggest a theoretical framework that highlights the role of individual and community identities in a way that helps to support an understanding of the challenges to recruitment and participation of underrepresented populations in research.

Measurement and Generalizability Bias.

Current practices in big data and open science often focus solely on measurements of internal validity, strict experimental control, and direct replicability to make inferences about generalizability that oftentimes ignore the importance of cultural context and sociodemographic factors (Hruschka et al., 2018). Attempts to quell the replicability crisis often target direct replication but lack emphasis on generalizability. When findings with limited generalizability are applied outside of academia, they can result in biased news reporting, policies, and diagnoses. Attention to the cultural context and accompanying lived experiences can prevent overgeneralization.

Beyond a meaningful sampling strategy, the appropriateness of the measure used to assess the lived experiences of the child being tested is critical to good internal and external validity. For example, cognitive testing using iPads, of a child who has not been exposed to interactive digital media, may result in biased results. Home questionnaires that ask about books in the home but not community sources of enrichment materials (e.g., public libraries, community centers etc.) may miss key structural resources of cognitive stimulation. Additionally, Bronfenbrenner (1974) described isolating a child in a room with a stranger as one of the most unnatural experiences and is likely to be received differently by children of different ages and from different communities. Experimental minutia such as testing context, specifically the family's overall level of comfort with the testing environment and researchers, can be a source of bias that needs to be considered prior to data collection.

To date, because most of the developmental literature has left out participants from racial/ethnic minority populations (Syed et al., 2018), many of the measures in our field are normed or primarily tested on a homogeneous sample of white middle income children. As such, the appropriateness of these measures to other American communities must be considered. Questions surrounding the fit of current standardized measures are especially concerning as a disproportionate number of families that experience early drug exposures and adversity flagged by HBCD might be non-white and living in under-resourced neighborhoods. In such an instance, one could imagine a scenario where variations in developmental trajectories emerge, not because of specific early life experiences, but because of the invalidity of the measures being used to assess development and performance. Indeed, a well-documented source of measurement bias is a lack of measurement invariance. Measurement invariance is the extent to which a psychometric test measures the same construct over repeated measurements or across different populations (Putnick & Bornstein, 2016; Wang et al., 2018). Measurement noninvariance is of particular concern for HBCD. Noninvariance across groups can occur in the form of differential item functioning (DIF). This is when psychometric test items assess subgroup (e.g., gender, language, education level, race, etc.) characteristics *in addition to* the construct of interest (Gibbons et al., 2011; Martinková et al., 2017). As a result, meaningful group differences cannot be confidently derived. DIF is a real risk for studies that use measures that have been normed for a specific subgroup (e.g., affluent White children) but are administered to other (e.g., lower-income and/or non-White children) groups (Lewis et al., 2012). Putnik and Bornstein (2016) provide an extensive review of measurement invariance issues in developmental science, as well as best practices for statistically testing for and reporting measurement invariance. Papers by Andrich and Hagquist also provide a comprehensive assessment of statistical methods for identifying specific types of DIF and methods for handling DIF when it occurs during data analysis (Andrich & Hagquist, 2012; Hagquist, 2019; Hagquist & Andrich, 2017). The time for considering these issues is at the outset of HBCD study design. Once data are collected and made widely available, these issues become shrouded in a potentially false sense of security for outside users.

In sum, measurement bias should be seriously considered by HBCD data collection sites and future data users. The primary goal of HBCD is to understand and document development across time as a function of age and environmental factors. By its nature, HBCD is designed to use the same measures over time and across multiple, disparate study sites. When choosing final measures for inclusion, it is critical to ensure true representative sampling when these measures were standardized, and to consider the similarity of the demographic makeup of the standardization sample and that of the current sample to be tested. As in the sampling bias issue, local community partnerships and liaisons can be a tremendous resource for understanding the lived experiences and challenges of communities and evaluating the appropriateness of measures and data interpretation in context.

Normativity Bias and Theory Building.

The normativity bias is when overgeneralized interpretations of data, derived from narrow sampling and measurement that lacks broad validity, specifically captures, and normalizes the experiences of dominant groups while 'othering' crucial aspects of non-dominant environmental realities. Deficit models contribute to biased narratives about development as they emphasize a reduced capacity of children growing up in non-dominant cultures, instead of interpreting their development within the contextual demands that are relevant to the child. For example, harshness and predictability are two key dimensions that shape development and can be used to characterize environmental experiences that positively and negatively impact an organism (Belsky et al., 2012; Ellis et al., 2009). Harshness constitutes the rate in which extrinsic factors cause disability and death; predictability is the degree of stochastic variation in environmental conditions (e.g., frequency of housing relocation) that regulate development (Ellis et al., 2009). High levels of harshness and unpredictability can result in stress, but stress can differentially shape any given child's development. Cognitive development that occurs to circumvent the stressor, or access additional resources are seen as adaptive for development. This stress-adapted trajectory is not accounted for in deficit models. Deficit models shape inaccurate narratives of often already underserved communities, which further marginalizes these children. The plausible alternative to the deficit model is that children from different populations experience differing degrees of harshness and predictability that bidirectionally shape their development.

When interventions are developed from theories formulated solely on the experiences of children from homes that do not experience varying degrees of adversity and stressors, these interventions are rendered ineffective. Even when researchers consider variation in experience, our theoretical frameworks may be rooted in an incomplete understanding of crucial spectra (i.e., harshness, unpredictability) as a vast majority of the research is based on the experiences of culturally, racially, and socioeconomically homogenous samples (Bornstein et al., 2013; Syed et al., 2018). Active changes in the form of theory building are necessary but require accurate and consistent conceptualizations of children's lived experiences. We urge scientists to develop an "interpretative power" or the ability to appreciate and account for the value of an individuals' experiences and behaviors within the scope of their cultural context (Brady et al., 2018). This interpretative power approach is best demonstrated in critical race theory, which emphasizes the role of identity and experience (see Delgado & Stefancic, 2017). The next section will introduce an interdisciplinary approach to developmental science interpretation that accounts for an individual's cultural context as well as its accompanying environmental exposures.

Towards A More Equitable and Accurate Developmental Neuroscience

Reducing bias in HBCD will not be an easy undertaking but is a worthwhile and attainable goal. This will require an unlearning of prior practices and an ongoing committed effort to learn alternative strategies. There will most certainly be discomfort or defensiveness of past approaches, but a genuine commitment towards equitable developmental neuroscience research within the large-scale HBCD study could guide the rest of the field forward. Ultimately, interpreting our research findings requires an interdisciplinary approach, interweaving observations and discoveries by developmental scientists, epidemiologists, economists, and critical race scholars. A holistic, interdisciplinary approach consists of an ecological account of development that considers the demands of a child's environment; a consideration of economic policy that emphasizes how individuals fare within the structures and systems we have constructed in American society; and a critical race theory approach that considers the compounding influence of an individual's identities within these socially constructed environments.

Developmental Systems Approach.

Ecological accounts of development conceptualize infants, children, and adults as distinct organisms that occupy different ecological niches, each with unique environmental demands that the organism must adapt to (Bronfenbrenner, 1974; Werchan & Amso, 2017). This ecological view asks the critical question: what purpose does a variation in development serve the organism at its current point in ontogeny and within the context in which it is developing? The aforementioned deficit models frame early life adversity as disadvantageous for development and capture the life-long "consequences" to development that come with chronic early-life stress (Rifkin-Graboi et al., 2021). While it is understood that disadvantage compounds over time (Hackman et al., 2015), longitudinal work suggests that context plays a role in how stress-developed cognitive functions perform (Nederhof et al., 2014; Nederhof & Schmidt, 2012; Young et al., 2018). Research that takes an alternative approach reframes the experience of adversity and identifies circumstances in which individual differences in brain structure and cognitive functioning may be adaptive to the well-being of the child at the time of the adversity (Frankenhuis & Nettle, 2020). Note, we are in no way suggesting that adversity is ever good. Rather, we are suggesting that one can acknowledge both that early life stressors and adversities must be alleviated, and simultaneously not erroneously judge development *in response* to meeting the needs of survival in difficult contexts as itself "not good". Instead, we must apply an unbiased approach to interpreting data on how children develop *in response* to stress and adversity.

Consider a data set that yields a positive association between socioeconomic status and executive functions. There are three possible explanations that are not mutually exclusive which directly shape our understanding of development. First, living in a low-income household is not as good for brain development as is living in a high-income household. This is inherently a deficit model interpretation. Alternatively, children living in a low-income household are indeed the norm (if we could measure every child on the planet), and children living in high income American households are pulling away from that norm (Amso et al., 2019). Third, brain development is adaptation to one's own lived experience. For high income children, this may be an abundance of conventionally enriching opportunities and activities that engages cortical resources. Children in under-resourced communities are adapting to fewer of these traditionally defined enriching experiences and their cortical surface area proportionally reflects their experience. This does not imply that their brains are damaged or likely to have different long-term trajectories. Fundamentally, these children have equally "good" brain development, meaning it is adapted to the challenges faced in their current environments (Amso, 2020)

and at an individual level, any one child, if offered a series of traditionally enriching community programs, would adapt accordingly. Currently, the field's reliance on normative ideology obstructs our ability to accurately interpret these adaptations. Put bluntly, the tacit, scientifically questionable assumption is that normal and good can be defined by data from predominantly white wealthy American children. This assumption does not come from science, but from a history of postcolonial systems and structures that define who and what is valuable in American society.

Returning to our example, bias occurs when the scientific community is bound to interpret the adaptation, possibly of the higher income group, as a deficit of the lower income group. This result of course fits and perpetuates established stereotypes. On the other hand, it is less common to interpret better socioemotional regulation in children with a history of poor maternal caregiving sensitivity as an advantage. However, data have shown better performance on memory tasks with emotional content in children with prior insensitive caregiving experiences (Rifkin-Graboi et al., 2021). Like the high-income children adapting to enrichment opportunities and performing better on executive functions tasks, it seems that children who have experienced poor maternal sensitivity are better at managing socioemotional information in memory. Certainly, the experience in which this developed is not an advantage, but how the brain adapts to it is as it should be on the measured timescale.

Our point is precisely that neither the high income/executive functions nor the poorer caregiving/emotional memory child has clearly better or worse brain development. Children with different lived experiences cannot be directly compared to each other without employing normativity bias. Engaging in these practices results in missed opportunities to understand limitations or adaptations in the face of unexplored environmental experiences (e.g., growing up in a place where attending school is not the norm), alternative types of environmental enrichment (e.g., taking on "adult" responsibilities at a younger age), and context relevant cognition based on environmental demands (e.g., exhibiting atypical abilities to plan and represent goals at a cognitively "young" age). Outside the field of developmental neuroscience exists critical work that further strengthens the interpretative power of developmental trajectories. For this, we can look to the fields of economics and critical race studies.

Interdisciplinary Approach: Looking Outside the Child and Household.

We can further build our interpretive power through incorporating work from the field of economics for example. Research within economics highlights a possible blind spot within developmental neuroscience. When thinking about how experience impacts the brain, developmental neuroscience has generally over-emphasized the role of family or household socioeconomic status (SES). As a result, it often excludes the larger impact of the ecological context of neighborhood or community (Amso & Lynn, 2017).

For example, findings from the Opportunity Insights Project (Chetty & Hendren, 2018) and the ABCD study (Taylor et al., 2020) suggest that ecological contexts outside of the child's household play a significant role in immediate and lifelong outcomes. Taylor et al. (2020) found that neighborhood poverty is associated with variation in prefrontal and hippocampal volumes. From the available ABCD data, we cannot know if these effects reflect maturation or adaptation, or whether they are fixed or reversible. However, this is a critical question and one that must be addressed by HBCD. What developmental process do our findings reflect? Chetty, Hendren, and colleagues found that children who grew up in high opportunity neighborhoods, regardless of income, had intergenerational mobility and higher adult income *in proportion with the time children spent in these areas* (Chetty et al., 2016; Chetty

& Hendren, 2018). They found that provided time in a better neighborhood, children adapt to the enriching context. When taken together with the economics data, the interpretation of the Taylor et al. (2020) findings on variation in prefrontal and hippocampal cortical volumes may best be interpreted as an adaptation to neighborhood poverty that is reversible, rather than a deficit in brain development that is fixed at an early age (Amso, 2020).

A focus on interpretive power requires that the HBCD study further explore how community resources, social policies, criminal justice policies, and neighborhood opportunities shape cognitive development independent of and in concert with the household context or specific pre- and postnatal exposures. For example, there are data on individuals who are considered to have "beaten the odds" of adversity or are resilient to their circumstances (Ellis et al., 2017; Masten, 2014). However, we argue that the odds themselves differ based on the socio-political identities of the individual. These odds can be further understood through the lens of intersectionality within critical race studies. Intersectionality is a term coined by Kimberlé Crenshaw (Crenshaw, 1991) and is an exposition of the idea that each person carries several political and social identities (e.g., race, gender), that these identities intersect, and that they can be a powerful force in creating opportunities for both oppression and advantage. Returning to the discussion above, the effects of neighborhoods on intergenerational mobility are not uniform but are moderated by race and gender. Economists have found that Black males do not get the same benefit of high opportunity neighborhoods as other groups, likely because of the contextual influences of systemic racism (Chetty, Hendren, Kline, & Saez, 2014).

Crenshaw's intersectional framework further argues that our identities compound and are both shaped by and actively shape our experiences. In scientific practice, however, we often treat these identities as nuisance variables. For example, exposures that are often experienced by children from marginalized communities are inextricably associated with racism and socioeconomic inequality (Amso & Lynn, 2017; Frankenhuis & Nettle, 2020). These intersections need to be clearly reflected in HBCD research and during theory building. We challenge scientists to not only understand these socio-political identities as more than category labels, but to undertake the challenge of understanding how they collaboratively shape development.

To achieve this goal, scientists will need to analyze how the challenges of opioid exposure, for example, are not isolated, but also how socio-political identities like race or residential status (e.g., residing in a rural vs more urban area) further compound and complicate this particular childhood stressor/adversity. An understanding of how structural systems differentially offer treatment vs. incarceration based on socio-political identities for opioid use is vital to comprehending the relationship between prenatal opioid use and later child outcomes. This permits differential interpretation of data from a predominantly White community where parents battling opioid addiction are offered treatment, whereas parents in neighboring predominantly Black communities may suffer prison sentences that deprive children of their caregivers (Hansen et al., 2020; James & Jordan, 2018; Saloner & Cook, 2013; Santoro & Santoro, 2018). When we integrate these realities with our established understanding of the significance of developmental mechanisms like maternal/parental modulation, we can develop a robust picture of the developmental demands facing children in this country. Instead of "othering" children from jurisdictions that do not offer recovery services, we can seek to understand how those children adapt to the conditions they are reared in. Examining the multiplicity of lived experiences permits the development of theory that can account for how these conditions can shape children.

Summary of considerations and recommendations

Large-scale studies such as ABCD and the upcoming HBCD study provide a unique opportunity to glean insight into conditions and factors that shape child development. Here, we highlight methodological checkpoints for researchers to reflect on to reduce bias at all phases of these studies.

During the initial phases of protocol development and participant recruitment, sampling bias and measurement bias should be of concern to researchers. In both cases, incorporating community members as liaisons, and even as research team members, offers an opportunity to: (a) partner with local communities, (b) fully understand their lived experiences, (c) consider how to maximize trust, (d) assess sampling and measurement appropriateness, and (e) make appropriate adjustments to protocols. The idea that local sites may alter testing details in a way that impacts the ability to coalesce data across all sites may at first blush seem like an immediate red flag in study design. We counteract this concern with two points. First, the study may already suffer from several biases for reasons stated throughout this manuscript. Second, there are solutions to this problem that include the possibility of the HBCD governing body forming a panel to consider protocol adjustment requests by local sites. The panel can make scientifically informed judgements about whether the requested shift is justified and appropriate.

During data analysis and interpretation, researchers should be mindful about biases impacting measurement and theory building. We note that initially avoiding sampling and measurement bias highly reduces the probability of erroneous and de-contextualized data interpretations. In order to further combat these issues and promote a more equitable and inclusive developmental science, we encourage researchers to look for measurement invariance and embrace a developmental systems and interdisciplinary approach. We understand this requires time and energy for scholarship that can be outside of some HBCD members' specific scientific domains. Inclusion of an external ethics board as well as an external committee composed of relevant social and cultural scientists from historically excluded groups will go a long way in ensuring that culturally relevant and data-appropriate measurements are used, and nuanced interpretations are made.

Conclusions

The purpose of this article was to highlight how standard practices within the field of developmental science can contribute to de-contextualized and stereotypical narratives about the development of children from under-resourced neighborhoods. As the goal of the proposed HBCD study is to better understand "normative" trajectories of brain development and investigate how drug exposures and environmental adversity impact neurocognitive development, we argue that it is imperative that neuroscientists involved in these, and other large-scale studies examine their own biases and engagement in these standard research practices. Careful considerations must be taken when interpreting results of families from marginalized communities. The NIH's request for application (RFA) for the HBCD study indicates that "*A deep nuanced understanding of factors that affect a child's health, brain and behavioral development is expected to emerge from this study, which is an essential first step toward designing policies and interventions that promote well-being and resilience in all children.*" In order to meet this goal, an intersectional framework must be embraced to adequately account for the unique experiences of children growing up in complex environments whose experiences are further complicated by their socio-political identities, and their accompanying systemic oppression. Not doing so will contribute to invalid science, an imprecise understanding of development, and will ultimately lead to ineffective policies that do not adequately address the needs of families across the country.

11

References

Amso, D. (2020). Neighborhood Poverty and Brain Development. *JAMA Network Open*, *3*(11), e2024139. https://doi.org/10.1001/jamanetworkopen.2020.24139

Amso, D., & Lynn, A. (2017). *Distinctive Mechanisms of Adversity and Socioeconomic Inequality in Child Development : A Review and Recommendations for Evidence-Based Policy*. https://doi.org/10.1177/2372732217721933

Amso, D., Salhi, C., & Badre, D. (2019). The relationship between cognitive enrichment and cognitive control: A systematic investigation of environmental influences on development through socioeconomic status. *Developmental Psychobiology*, *March*, 1–20. https://doi.org/10.1002/dev.21794

Andrich, D., & Hagquist, C. (2012). Real and Artificial Differential Item Functioning. *Journal of Educational and Behavioral Statistics*, *37*(3), 387–416. https://doi.org/10.3102/1076998611411913

Belsky, J., Schlomer, G. L., & Ellis, B. J. (2012). Beyond cumulative risk: Distinguishing harshness and unpredictability as determinants of parenting and early life history strategy. *Developmental Psychology*, *48*(3), 662–673. https://doi.org/10.1037/a0024454

Bonevski, B., Randell, M., Paul, C., Chapman, K., Twyman, L., Bryant, J., Brozek, I., & Hughes, C. (2014). Reaching the hard-to-reach: A systematic review of strategies for improving health and medical research with socially disadvantaged groups. *BMC Medical Research Methodology*, *14*(1). https://doi.org/10.1186/1471-2288-14-42

Bornstein, M., Jager, J., & Putnick, D. (2013). Sampling in developmental science: Situations, shortcomings, solutions, and standards. In *Developmental Review* (Vol. 33, Issue 4, pp. 357–370). https://doi.org/10.1016/j.dr.2013.08.003

Brady, L. M., Fryberg, S. A., & Shoda, Y. (2018). Expanding the interpretive power of psychological science by attending to culture. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(45), 11406–11413. https://doi.org/10.1073/pnas.1803526115

Bronfenbrenner, U. (1974). Developmental Research, Public Policy, and the Ecology of Childhood. *Child Development*, *45*(1), 1–5. https://about.jstor.org/terms

Burger, J., & Gochfeld, M. (2011). *Conceptual environmental justice model for evaluating chemical pathways of exposure in low-income, minority, Native American, and other unique exposure populations*. *101*, 64–73. https://doi.org/10.2105/AJPH

Carlson, M. D. A., & Morrison, R. S. (2009). Study design, precision, and validity in observational studies. *Journal of Palliative Medicine*, *12*(1), 77–82. https://doi.org/10.1089/jpm.2008.9690

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics*, *129*(4), 1553–1623.

Chetty, R., & Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility II: County-level estimates. *Quarterly Journal of Economics*, *133*(3), 1163–1228. https://doi.org/10.1093/QJE/QJY006

Chetty, R., Hendren, N., & Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. In *American Economic Review* (Vol. 106, Issue 4, pp. 855–902). American Economic Association. https://doi.org/10.1257/aer.20150572

Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. In *Source: Stanford Law Review* (Vol. 43, Issue 6).

Delgado, R., & Stefancic, J. (2017). *Critical race theory : an introduction*. New York University Press.

Ellis, B. J., Bianchi, J. M., Griskevicius, V., & Frankenhuis, W. E. (2017). Beyond Risk and Protective Factors: An Adaptation-Based Approach to Resilience. *Perspectives on Psychological Science*, *12*(4), 561–587. https://doi.org/10.1177/1745691617693054

Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & Schlomer, G. L. (2009). Fundamental dimensions of environmental risk: The impact of harsh versus unpredictable environments on the evolution and development of life history strategies. In *Human Nature* (Vol. 20, Issue 2). https://doi.org/10.1007/s12110-009-9063-7

Flora, D. B. (2020). Thinking about effect sizes: From the replication crisis to a cumulative psychological science. *Canadian Psychology*, *61*(4), 318–330. https://doi.org/10.1037/cap0000218

Frankenhuis, W. E., & Nettle, D. (2020). The Strengths of People in Poverty. *Current Directions in Psychological Science*, *29*(1), 16–21. https://doi.org/10.1177/0963721419881154

French, B. H., Lewis, J. A., Mosley, D. V., Adames, H. Y., Chavez-Dueñas, N. Y., Chen, G. A., & Neville, H. A. (2020). Toward a Psychological Framework of Radical Healing in Communities of Color. *Counseling Psychologist*, *48*(1), 14–46. https://doi.org/10.1177/0011000019843506

Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R. Z., Heeringa, S., Jernigan, T., Potter, A., Thompson, W., & Zahs, D. (2018). Recruiting the ABCD sample: Design considerations and procedures. *Developmental Cognitive Neuroscience*, *32*, 16–22. https://doi.org/10.1016/j.dcn.2018.04.004

Gibbons, L., Crane, P. K., Mehta, K. M., Pedraza, O., Tang, Y., Manly, J. J., Narasimhalu, K., Teresi, J., Jones, R. N., & Mungas, D. (2011). Multiple, correlated covariates associated with differential item functioning (DIF): Accounting for language DIF when education levels differ across languages. *Ageing Research*, *2*(1), 4. https://doi.org/10.4081/ar.2011.e4

Hackman, D. A., Gallop, R., Evans, G. W., & Farah, M. J. (2015). Socioeconomic status and executive function: Developmental trajectories and mediation. *Developmental Science*, *18*(5), 686–702. https://doi.org/10.1111/desc.12246

Hagquist, C. (2019). Explaining differential item functioning focusing on the crucial role of external

information - an example from the measurement of adolescent mental health. *BMC Medical Research Methodology*, *19*(1), 185. https://doi.org/10.1186/s12874-019-0828-3

Hagquist, C., & Andrich, D. (2017). Recent advances in analysis of differential item functioning in health research using the Rasch model. *Health and Quality of Life Outcomes*, *15*(1). https://doi.org/10.1186/s12955-017-0755-0

Hansen, H., Parker, C., & Netherland, J. (2020). Race as a Ghost Variable in (White) Opioid Research. *Science Technology and Human Values*, *45*(5), 848–876. https://doi.org/10.1177/0162243920912812

Hruschka, D. J., Medin, D. L., Rogoff, B., & Henrich, J. (2018). Pressing questions in the study of psychological and behavioral diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(45), 11366–11368. https://doi.org/10.1073/pnas.1814733115

James, K., & Jordan, A. (2018). The opioid crisis in black communities. *Journal of Law, Medicine and Ethics*, *46*(2), 404–421. https://doi.org/10.1177/1073110518782949

Johnson, M. H., Jones, E. J. H., & Gliga, T. (2015). Brain adaptation and alternative developmental trajectories. *Development and Psychopathology*, *27*(2), 425–442. https://doi.org/10.1017/S0954579415000073

Lewis, T. T., Yang, F. M., Jacobs, E. A., & Fitchett, G. (2012). Racial/ethnic differences in responses to the everyday discrimination scale: A differential item functioning analysis. *American Journal of Epidemiology*, *175*(5), 391–401. https://doi.org/10.1093/aje/kwr287

Martinková, P., Drabinová, A., Liaw, Y. L., Sanders, E. A., McFarland, J. L., & Price, R. M. (2017). Checking equity: Why differential item functioning analysis should be a routine part of developing conceptual assessments. *CBE Life Sciences Education*, *16*(2). https://doi.org/10.1187/cbe.16-10-0307

Masten, A. S. (2014). Global Perspectives on Resilience in Children and Youth. *Child Development*, *85*(1), 6–20. https://doi.org/10.1111/cdev.12205

Nederhof, E., Ormel, J., & Oldehinkel, A. J. (2014). *Mismatch or Cumulative Stress : The Pathway to Depression Is Conditional on Attention Style*. https://doi.org/10.1177/0956797613513473

Nederhof, E., & Schmidt, M. V. (2012). Mismatch or cumulative stress: Toward an integrated hypothesis of programming effects. *Physiology and Behavior*, *106*(5), 701–706. https://doi.org/10.1016/j.physbeh.2011.12.008

Pullum, G. K. (1999). African American Vernacular English is not standard English with mistakes. In *he workings of language: From prescriptions to perspectives* (pp. 59–66).

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. In *Developmental Review* (Vol. 41, pp. 71–90). Mosby Inc. https://doi.org/10.1016/j.dr.2016.06.004

Rifkin-Graboi, A., Goh, S. K.-Y., Chong, H. J., Tsotsi, S., Sim, L. W., Tan, K. H., Chong, Y. S., & Meaney, M. J. (2021). Caregiving adversity during infancy and preschool cognitive function: adaptations to context? *Journal of Developmental Origins of Health and Disease*, 1–12. https://doi.org/10.1017/S2040174420001348

Saloner, B., & Cook, B. L. (2013). Blacks and hispanics are less likely than whites to complete addiction treatment, largely due to socioeconomic factors. *Health Affairs*, *32*(1), 135–145. https://doi.org/10.1377/hlthaff.2011.0983

Santoro, T. N., & Santoro, J. D. (2018). Racial Bias in the US Opioid Epidemic: A Review of the History of Systemic Bias and Implications for Care. *Cureus*. https://doi.org/10.7759/cureus.3733

Scharf, D. P., Mathews, K. J., Jackson, P., Hofsuemmer, J., Martin, E., & Edwards, D. (2010). More than Tuskegee: Understanding mistrust about research participation. *Journal of Health Care for the Poor and Underserved*, *21*(3), 879–897. https://doi.org/10.1353/hpu.0.0323

Serdar, C. C., Cihan, M., Yücel, D., & Serdar, M. A. (2021). Sample size, power and effect size revisited: Simplified and practical approachin pre-clinical, clinical and laboratory studies. *Biochemia Medica*, *31*(1), 1–27. https://doi.org/10.11613/BM.2021.010502

Simmons, C., Conley, M. I., Gee, D. G., Baskin-Sommers, A., Barch, D. M., Hoffman, E. A., Huber, R. S., Iacono, W. G., Nagel, B. J., Palmer, C. E., Sheth, C. S., Sowell, E. R., Thompson, W. K., & Casey, B. J. (2021). Responsible Use of Open-Access Developmental Data: The Adolescent Brain Cognitive Development (ABCD) Study. *Psychological Science*, *32*(6), 866–870. https://doi.org/10.1177/09567976211003564

Sullivan, G. M., & Feinn, R. (2012). Using Effect Size—or Why the P Value Is Not Enough . *Journal of Graduate Medical Education*, *4*(3), 279–282. https://doi.org/10.4300/jgme-d-12-00156.1

Syed, M., Santos, C., Yoo, H. C., & Juang, L. P. (2018). Invisibility of racial/ethnic minorities in developmental science: Implications for research and institutional practices. *American Psychologist*, *73*(6), 812–826. https://doi.org/10.1037/amp0000294

Taylor, R. L., Cooper, S. R., Jackson, J. J., & Barch, D. M. (2020). Assessment of Neighborhood Poverty, Cognitive Function, and Prefrontal and Hippocampal Volumes in Children. *JAMA Network Open*, *3*(11), e2023774. https://doi.org/10.1001/jamanetworkopen.2020.23774

Thiese, M. S. (2014). Observational and interventional study design types; an overview. *Biochemia Medica*, *24*(2), 199–210. https://doi.org/10.11613/BM.2014.022

Wang, S., Chen, C. C., Dai, C. L., & Richardson, G. B. (2018). A Call for, and Beginner's Guide to, Measurement Invariance Testing in Evolutionary Psychology. *Evolutionary Psychological Science*, *4*(2), 166–178. https://doi.org/10.1007/s40806-017-0125-5

Washington, H. A. (2006). *Medical apartheid: The dark history of medical experimentation on Black Americans from colonial times to the present*. Doubleday Books.

Werchan, D. M., & Amso, D. (2017). A novel ecological account of prefrontal cortex functional

development. *Psychological Review*, *124*(6), 720–739. https://doi.org/10.1037/rev0000078

Williams, D. R., Priest, N., & Anderson, N. B. (2016). Understanding associations among race, socioeconomic status, and health: Patterns and prospects. In *Health Psychology* (Vol. 35, Issue 4, pp. 407–411). American Psychological Association Inc. https://doi.org/10.1037/hea0000242

Young, E. S., Griskevicius, V., Simpson, J. A., & Waters, T. E. A. (2018). Can an Unpredictable Childhood Environment Enhance Working Memory? Testing the Sensitized-Specialization Hypothesis. *Journal of Personality and Social Psychology*, *February*. https://doi.org/10.1037/pspi0000124

Zuo, X. N., Xu, T., & Milham, M. P. (2019). Harnessing reliability for neuroscience research. In *Nature Human Behaviour*. Nature Research. https://doi.org/10.1038/s41562-019-0655-x

Declaration of Competing Interest

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: